Close Counterfactuals and Almost Doing the Impossible

Tiffany Doan, Stephanie Denison, and Ori Friedman

Department of Psychology, University of Waterloo

Date resubmitted: April 17, 2023

Author Note

Correspondence concerning this article should be addressed to Tiffany Doan, Department of Psychology, University of Waterloo, 200 University Avenue W, Waterloo, Ontario, Canada N2L 3G1. E-mail: t3doan@uwaterloo.ca

Open Practices Statement: The preregistrations, stimuli, and data from all experiments are available on the Open Science Framework, https://osf.io/vyd7s/?view_only=671f7b63a1a144c68e231f66af7a456b.

Abstract

Can we feel that an unrealized outcome nearly happened if it was never possible in the first place? People often consider counterfactual events that did not happen, and some counterfactuals seem so close to reality that people say they "almost" or "easily could have" happened. Across four preregistered experiments (total N = 1228), we investigated how judgments of counterfactual closeness depend on possibility, and whether this varies across two kinds of close counterfactuals. In judging whether outcomes almost happened, participants were more strongly impacted by possibility than by incremental manipulations of probability. In contrast, when judging whether outcomes easily could have happened, participants treated the distinction between impossible and possible like any other variation in probability. Both kinds of judgments were also impacted by propensity, though these effects were comparatively small. Together, these findings reveal novel differences between the two kinds of close counterfactuals and suggest that while possibility is privileged when judging what almost happened, probability is the focus when judging what easily could have happened.

Keywords: counterfactuals, possibility, propensity, probability, cognition

Close Counterfactuals and Almost Doing the Impossible

In allowing us to think about how things could be different than they really are, counterfactuals cover a lot of ground. Some counterfactuals are fantastical and outlandish. For example, you can speculate about what life would be like if you were a bird or if time travel were possible. Other counterfactuals are more plausible, as when you wonder where you would be today if you had taken a different job, or conclude that your presentation would have been better if you'd gone to bed on time last night. These counterfactuals are invoked in conditional statements about how things might have gone in another world (e.g., "If I'd gone to bed on time, my presentation would have been better"). However, another group of "close" counterfactuals feel so plausible and relevant for our lives that we describe them as though they were unrealized features of the actual world (Kahneman & Varey, 1990). For instance, we say a narrowly missed opportunity *almost* or *easily could have happened* (e.g., "it almost fell off the table"; Beck & Guthrie, 2011).

What makes some counterfactuals seem close? Kahneman and Varey (1990) suggested that closeness depends on *propensity*—it arises when preceding events were leading towards the counterfactual outcome (also see Covey & Zhang, 2014). To illustrate, we might say a golfer almost hit a hole-in-one if her ball was headed straight for the hole but then happened to stop just short of it. One interesting feature of this account is that it sharply distinguishes between prior probability and possibility. Whether the golfer almost hit a hole-in-one doesn't really depend on her prior odds of doing so. Even for professional golfers, the odds of a hole-in-one are vanishingly low ("Odds of a hole in one", 2021). Nonetheless, the propensity account suggests that counterfactual closeness requires possibility—it would be odd to assert that events were

leading to an impossible outcome. So, hitting a hole-in-one could not be counterfactually close if this outcome were made impossible, as might happen if someone filled the hole with cement.[1]

In this paper, we examine the importance of possibility for judgments of counterfactual closeness. Previous research on counterfactual closeness has not done this. Instead, research has investigated how closeness is affected by different instantiations of distance (e.g., Meyers-Levy & Maheswaran, 1992; Ong et al., 2015, 2021; Teigen, 1996). For example, people think it feels worse to miss a flight by 5 minutes than by 30 minutes (temporal distance; Kahneman & Tversky, 1982) and are more likely to say a dart almost hit a bullseye if it landed close rather than far from it (spatial distance; Gerstenberg & Tenenbaum, 2016). Prior work has also looked at how closeness is affected by prior probability and by people's expectations (e.g., Doan et al., 2021; Gerstenberg & Tenenbaum, 2016; Kahneman & Varey, 1990). Beyond this, prior research has examined how close counterfactuals affect people's emotions (e.g., Kühberger et al., 2011; Markman & Tetlock, 2000; Medvec & Savitsky, 1997; Mellers et al., 1997; Sweeny & Vohs, 2012), and how people judge the plausibility of counterfactual outcomes by assessing the similarity between reality and counterfactual worlds (e.g., De Brigard et al. 2021).

The point of departure for our investigation is recent evidence that there are at least two kinds of close counterfactuals. In introducing close counterfactuals, Kahneman and Varey (1990) focused exclusively on judgments of whether outcomes almost happened. But closeness can also be expressed by saying an outcome easily could have happened (e.g., Levy, 2015; Teigen, 1995, 1996; also see Chapter 2 in Kraft, 2012), and people draw different inferences about these two

[1] Strictly, hitting a hole-in-one into a hole filled with cement might not be impossible—while this outcome violates our physical laws, for all we know, these laws could change. Even so, most people may not see things this way, and may instead see violations of physical and biological laws as impossible. For example, people mostly deny that a person could walk through a wall or make a car vanish into thin air (e.g., Goulding & Friedman, in press; Philips & Cushman, 2017; Shtulman & Carey, 2007; Shtulman, 2009). These findings are not decisive though. For example, respondents could have interpreted questions about possibility as referring to events occurring in local contexts (i.e., where familiar physical and biological laws operate).

forms of counterfactual closeness (Doan et al., 2023). Besides depending on propensity, judgments that a counterfactual outcome almost happened (henceforth Almost judgments) depend on its perceived distance from the actual outcome (Beck & Guthrie, 2011; Gerstenberg & Tenenbaum, 2016; Kahneman & Varey, 1990). By contrast, judgments that an outcome easily could have happened (henceforth Easily judgments) depend more on its prior probability. Consider a luck-based contest where prizes are awarded when a lottery machine dispenses 6 red balls and no blue ones. If a contestant starts with good odds but her outcome is distant from what is needed to win (the balls in her machine are mostly red, but she gets many blue ones), participants affirm she easily could have won, but are less likely to agree she almost won (Doan et al., 2023). Conversely, if she starts with terrible odds but nonetheless comes close to winning (the balls in her machine are mostly blue, but she gets five red ones), people affirm she almost won, but deny she easily could have.

The question of how possibility contributes to counterfactual closeness, then, may warrant distinct answers for the two kinds of close counterfactuals. For instance, the distinction between possible versus impossible could be *especially* important for Almost judgments (i.e., as implied by the proposal that this kind of closeness depends on propensity; Kahneman & Varey, 1990). But the distinction between possible and impossible might not be privileged in Easily judgments. To illustrate: if an initially impossible outcome becomes increasingly likely, Almost judgments might be especially sensitive to the initial shift from impossible to possible, whereas Easily judgments might be similarly sensitive to all increases in probability. Along the same lines, the importance of propensity might also differ between these two kinds of close counterfactuals. Beyond revealing how possibility and propensity matter for judgments of

counterfactual closeness, the current investigation might also tell us more about how the two kinds of close counterfactuals differ from one another.

General Methods

We conducted four preregistered experiments. In each experiment, participants saw vignettes about contestants who tried to win prizes using scratch cards. The contestants always lost, and participants rated whether they almost won or easily could have won. Across the experiments, we manipulated the prior odds offered by the scratch cards and propensity—whether the contestants had initially seemed headed for a win.

Preregistrations, stimuli, and data from all experiments are available on the Open Science Framework, https://osf.io/vyd7s/?view_only=671f7b63a1a144c68e231f66af7a456b. Participants in all experiments were recruited using Cloud Research with the criteria that they were from the United States, had a HIT approval rating between 95-100%, and had completed at least 100 HITs. We tested different participants in each experiment and aimed to have at least 100-200 usable participants per between-subjects condition after exclusions. These sample sizes were based on previous studies using similar designs (Klusowski et al., 2021; Malter et al., 2021; Ong et al., 2018). This research was approved by the [blinded].

After participants completed the main task in each experiment, they were asked two additional multiple-choice comprehension questions to check that they were paying attention during the task. One asked which shapes they saw, and the other asked what judgment they made (i.e., the test question). Participants were excluded if they failed at least one comprehension question or if they neglected to answer any test questions. Participants were also asked about their age and gender.

The analyses in all experiments used generalized estimating equations models (GEE; linear; independent correlation matrix) and interactions were followed up with Bonferroni-corrected pairwise comparisons.

Experiment 1

Methods

Participants. The experiment was successfully completed by 206 adults (mean age = 37 years; 71 women, 133 men, 2 other/prefer not to say). An additional 22 participants were excluded for failing at least one of our two comprehension check questions at the end of the study.

Materials and procedure. Participants read about a contest that used scratch cards with eight covered squares. Under each square is one of three shapes: a star, clover, or heart. Each contestant can choose any four squares to scratch. If all four squares reveal the same shape, the contestant wins $1000. Following this information, participants were asked a multiple-choice comprehension question to confirm that they understood (i.e., "To win $1000, the contestants have to scratch __ identical shapes."). Participants who answered incorrectly were informed of this, were taken back to the page with the contest rules, and were asked the comprehension question again. If they answered incorrectly again, the story would repeat one more time. The task continued regardless of whether participants answered the comprehension question correctly the third time (no participants answered incorrectly three times).

In the main task, participants read about a series of contestants who lost the contest. For each contestant, participants first saw the four shapes the contestant chose (including the order in which they were scratched), and then the remaining shapes on the card. The experiment used a 2X2X2 design. Information about each of four contestants corresponded with cells yielded by

crossing two within-subjects factors; see Figure 1. One factor was propensity—whether the preceding events were leading towards a win. Low propensity contestants were not headed for a win—they lost because their second shape differed from the others. By contrast, high propensity contestants were initially headed towards a win. They scratched three matching shapes, and only lost because their fourth shape did not match. The other factor was the number of additional target shapes left on the scratch card (i.e., zero versus two target shapes remaining). When no target shapes remained, contestants could not choose another target shape, making it impossible for them to win with the shapes they had chosen. However, when two target shapes remained, contestants had a chance of choosing another target shape and winning.

After reading about each contestant, participants were asked a close counterfactual question about them losing. This test question was manipulated between-subjects. Participants either rated whether the contestant *almost* won (e.g., "Did Jamie almost win?") or *easily could have* won (e.g., "Could Jamie easily have won?") using a 7-point scale ranging from "Definitely No" (1) to "Definitely Yes" (7).

Participants also saw two extra trials presented with the four test trials to add variability to the outcomes and to make our manipulation less obvious to participants. In these trials, the contestants also lost, but had fewer matching shapes; see stimuli on OSF. Trial order was randomized across participants.

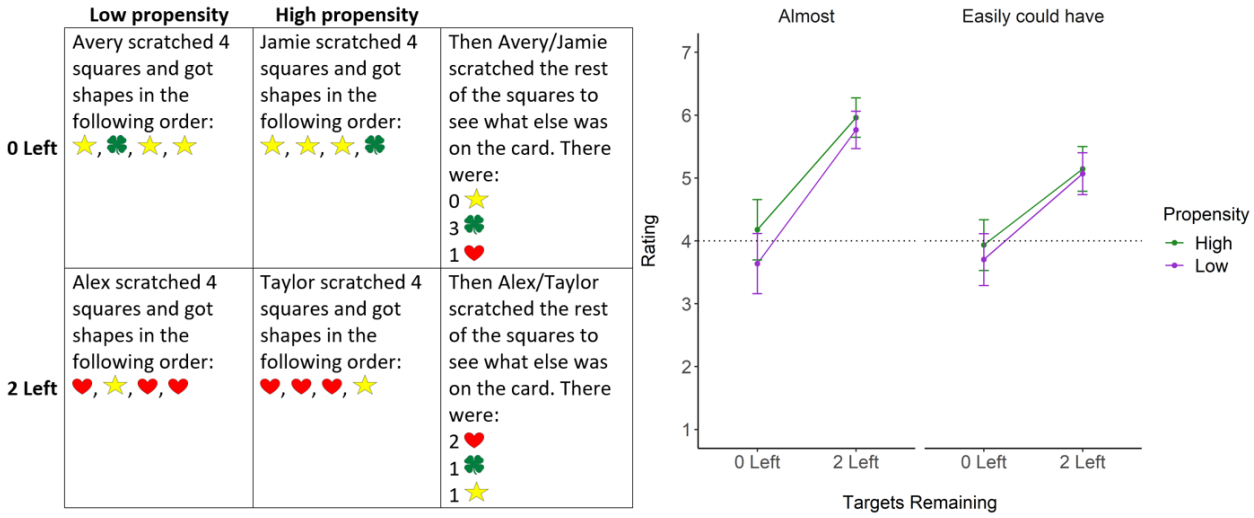

Figure 1. Stimuli of the test trials (left) and participants' responses to the close counterfactual questions (right) for Experiment 1. Vertical bars show 95% confidence intervals.

Results and Discussion

Figure 1 shows participants' mean ratings of counterfactual closeness. Ratings were higher for Almost than Easily judgments, when propensity was high rather than low, and when winning with the target was possible rather than impossible. Also, participants' sensitivity to possibility (i.e., higher ratings when two targets remained compared with when none remained) was greater for Almost than Easily ratings.

These observations were corroborated by a 2x2x2 GEE model with propensity (low, high) and the number of targets remaining (0, 2) as within-subjects factors and expression (Almost, Easily) as a between-subjects factor. All three main effects were significant: expression (Almost>Easily), *Wald* $X^2(1) = 4.16$, $p = .041$; propensity (high>low), *Wald* $X^2(1) = 13.00$, $p < .001$; and targets remaining (2>0), *Wald* $X^2(1) = 125.80$, $p < .001$. There was also an interaction between expression and the number of targets remaining, *Wald* $X^2(1) = 5.32$, $p = .021$; all other $ps \geq .139$. Overall, participants gave higher ratings when 2 targets remained than when none

remained, but this effect was larger for Almost than Easily judgments (Almost, $M_{difference}$ = 1.96, $p < .001$, CI95% = [1.35, 2.57]; Easily, $M_{difference}$ = 1.29, $p < .001$, CI95% = [0.83, 1.75]).[2]

These findings suggest the two kinds of close counterfactual judgments might be differentially sensitive to the distinction between impossible and possible outcomes. Alternatively, the pattern could indicate differing sensitivities to lower versus higher probability. We next tested between these possibilities using scenarios where the losing cards had either one or three additional targets remaining.

Experiment 2

Methods

Participants. The experiment was successfully completed by 218 adults (mean age = 40 years; 103 women, 112 men, 3 other/prefer not to say). An additional 14 participants were excluded for failing at least one of our two comprehension check questions at the end of the study.

Materials and procedure. The experiment was identical to Experiment 1, except each card had either one or three target shapes remaining (i.e., rather than zero or two); see Figure 2. For example, after scratching three heart shapes and one star shape, a contestant might then learn that there is one additional heart shape left on the card or three additional hearts left on the card.

This design manipulates probability without manipulating possibility because in both cases, winning with the target shape is possible. If Almost judgments do not continue to be more sensitive than Easily judgments to the difference between one and three target shapes remaining,

[2] In examining Figure 1, it might seem odd that Easily judgments were not at floor even when winning with the target shape was impossible. However, participants likely recognized that the contestant could have won with *other* shapes on their scratch card. Findings from one of the extra trials (included to add variability to the outcomes) support this interpretation. In that trial, winning was outright impossible—there was no way to scratch 4 matching shapes (i.e., the contestant scratched 2 hearts, 1 clover, and 1 star, and the remaining shapes were 1 heart, 2 clovers, and 1 star). Easily ratings were 1.50 on the 1-7 scale, and Almost ratings were 1.96. Experiment 2 also includes this trial, and there, Easily ratings were 1.54 and Almost ratings were 1.74.

this would suggest that the two kinds of close counterfactual judgments differ in their sensitivities between impossible and possible outcomes. However, if Almost judgments continue to be more sensitive than Easily judgments, this would suggest that the two judgments differ in their sensitivities to lower versus higher probability.

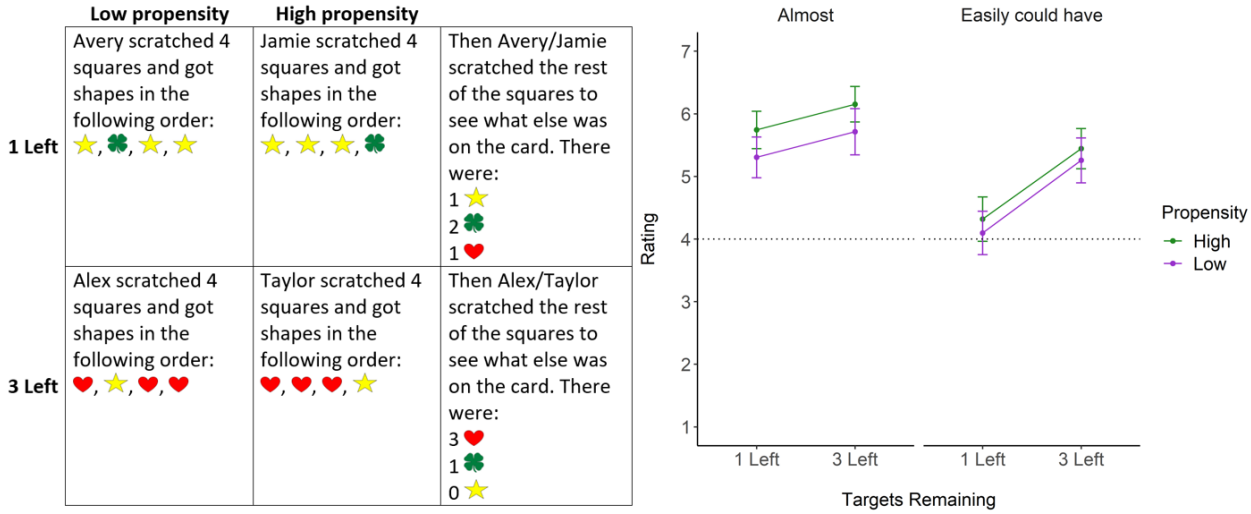

Figure 2. Stimuli of the test trials (left) and participants' responses to the close counterfactual questions (right) for Experiment 2. Vertical bars show 95% confidence intervals.

Results and Discussion

Figure 2 shows participants' mean ratings of counterfactual closeness. Ratings were higher for Almost judgments than Easily judgments, when propensity was high rather than low, and when three target shapes remained rather than one. Also, while participants gave higher ratings when more targets remained, this effect was larger for Easily than Almost ratings.

The observations are supported by a 2x2x2 GEE model with propensity (low, high) and the number of targets remaining (1, 3) as within-subjects factors, and expression (Almost, Easily) as a between-subjects factor. All three main effects were significant: expression (Almost>Easily), *Wald* $X^2(1) = 23.35$, $p < .001$; propensity (high>low), *Wald* $X^2(1) = 24.67$, $p < .001$; targets remaining (3>1), *Wald* $X^2(1) = 60.92$, $p < .001$. There was also an interaction

between expression and the number of targets remaining, *Wald* $X^2(1) = 13.57$, $p < .001$; all other ps $\geq .069$. The interaction resulted because the effect of number of targets remaining was larger for Easily than Almost ratings (Easily, $M_{difference} = 1.14$, $p < .001$, CI95% = [0.77, 1.51]; Almost, $M_{difference} = 0.41$, $p = .021$, CI95% = [0.04, 0.78]).

These findings are opposite to those from our first experiment. In Experiment 1, Almost judgments were *more* sensitive than Easily judgments to the difference between zero and two target shapes remaining. However, in this experiment, Almost judgments were *less* sensitive than Easily judgments to the difference between one and three target shapes remaining. What might explain this difference?

One answer is that Almost judgments are particularly sensitive to changes between impossible and possible situations, whereas Easily judgments are more sensitive to changes in probability. That is, Almost judgments may increase greatly when the likelihood of winning with the target shape changes from impossible (zero remaining) to possible (one or more remaining), but increases in smaller increments when winning is possible (e.g., one vs three remaining). Easily judgments may follow a different pattern in which ratings steadily increase as the likelihood of winning with the target shape increases (e.g., from zero to one to two remaining).

However, this explanation rests on informal comparisons across experiments. In the next experiment, we test this explanation directly by having zero, two, or four target shapes remaining on the scratch cards.

Experiment 3

Methods

Participants. The experiment was successfully completed by 397 adults (mean age = 41 years; 195 women, 201 men, 1 other/prefer not to say). An additional 35 participants were

excluded for failing our initial comprehension question three times or at least one of our two comprehension check questions at the end of the study.

Materials and procedure. The experiment was identical to the previous experiments, except scratch cards had 10 covered squares, and each card had either zero, two, or four target shapes remaining; see Figure 3.

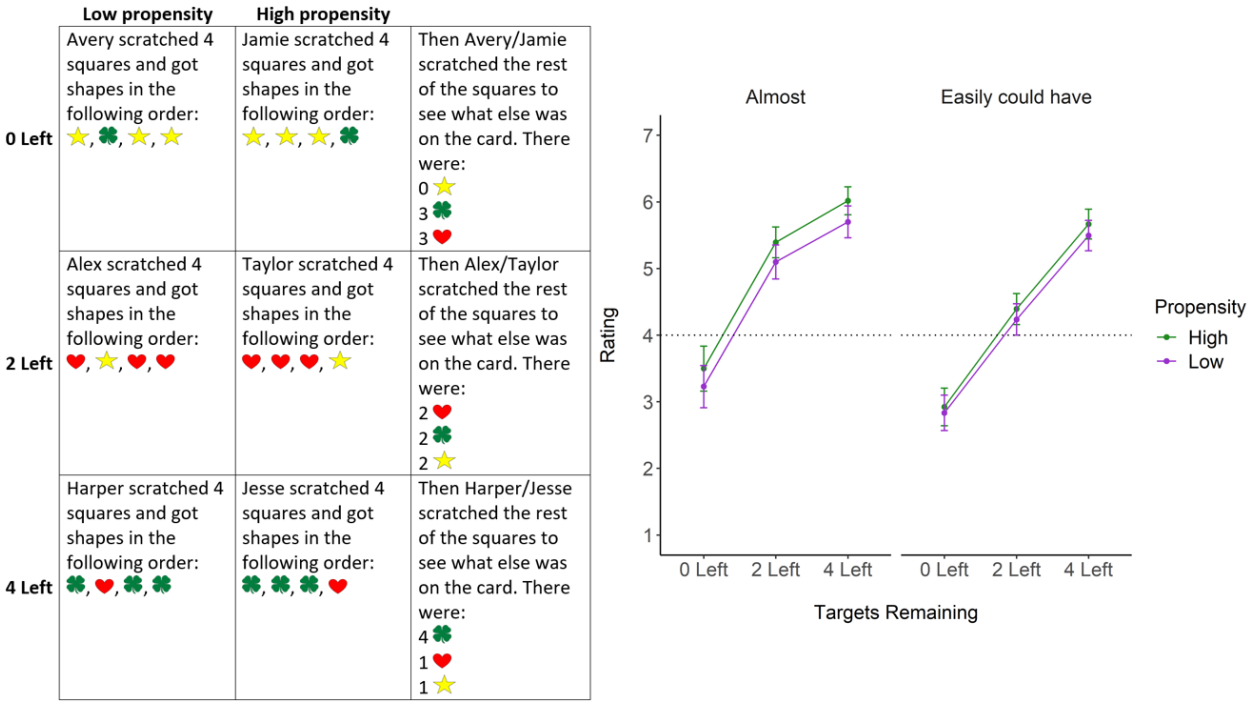

Figure 3. Stimuli of the test trials (left) and participants' responses to the close counterfactual questions (right) for Experiment 3. Vertical bars show 95% confidence intervals.

Results and Discussion

Figure 3 shows participants' mean ratings of counterfactual closeness. Ratings were higher for Almost than Easily judgments, when propensity was high rather than low, and when more target shapes remained. Also, Easily and Almost ratings were differentially affected by the number of targets remaining. With Almost judgments, participants sharply distinguished between situations where winning with the target was impossible versus possible (0 vs 2 remaining), but were comparatively less sensitive to further differences in possibility (2 vs 4 targets remaining).

By contrast, with Easily Judgments, participants did not show this discontinuity—instead, their ratings linearly increased as the probability of winning with the target increased.

These observations are supported by a 2x3x2 GEE model with propensity (low, high) and the number of targets remaining (0, 2, 4) as within-subjects factors, and expression (Almost, Easily) as a between-subjects factor. All three main effects were significant: expression (Almost>Easily), $Wald\ X^2(1) = 19.02$, $p < .001$; propensity (high>low), $Wald\ X^2(1) = 26.15$, $p < .001$; targets remaining (more>less), $Wald\ X^2(2) = 530.90$, $p < .001$. There was also an interaction between expression and the number of targets remaining, $Wald\ X^2(2) = 27.75$, $p < .001$; all other ps $\geq .067$.

The interaction resulted because the distinction between 0 and 2 targets remaining more strongly impacted Almost than Easily ratings (Almost, $M_{difference} = 1.88$, $p < .001$, CI95% = [1.39, 2.38]; Easily, $M_{difference} = 1.44$, $p < .001$, CI95% = [1.12, 1.76]), whereas the distinction between 2 and 4 remaining more strongly affected Easily than Almost ratings (Easily, $M_{difference} = 1.27$, $p < .001$, CI95% = [0.97, 1.56]; Almost, $M_{difference} = 0.61$, $p < .001$, CI95% = [0.36, 0.86]). Also, Almost ratings showed a greater increase when moving from 0 to 2 targets remaining than when moving from 2 to 4 remaining, $p < .001$, CI95% = [0.96, 1.59]. However, for Easily judgments, the increase in ratings from 0 to 2 remaining and from 2 to 4 remaining did not significantly differ, $p = .203$, CI95% = [-0.09, 0.43].

The findings support the idea that Almost judgments increase substantially when the likelihood of winning with the target shape changes from impossible to possible but only increase incrementally once winning with the target shape is possible. Easily judgments instead increase in a steady linear fashion as the number of targets remaining increase.

So far, we increased the number of target shapes remaining in increments of two (e.g., zero versus two and one versus three targets left). However, we did not examine the exact point at which the impossible changes to possible—the differences between zero shapes remaining and one shape remaining. The final experiment does this.

Experiment 4

Methods

Participants. The experiment was successfully completed by 407 adults (mean age = 42 years; 209 women, 194 men, 4 other/prefer not to say). An additional 25 participants were excluded for failing our initial comprehension question three times or at least one of our two comprehension check questions at the end of the study.

Materials and procedure. The experiment was similar to Experiment 3, except we added an additional extra trial and used a 9-point rating scale. We also did not manipulate propensity as it did not interact with other factors in previous experiments. Test trials showed contestants scratched three identical shapes and then a different shape on the fourth scratch. Each scratch card had either zero, one, two, or three target shapes remaining; see Figure 4.

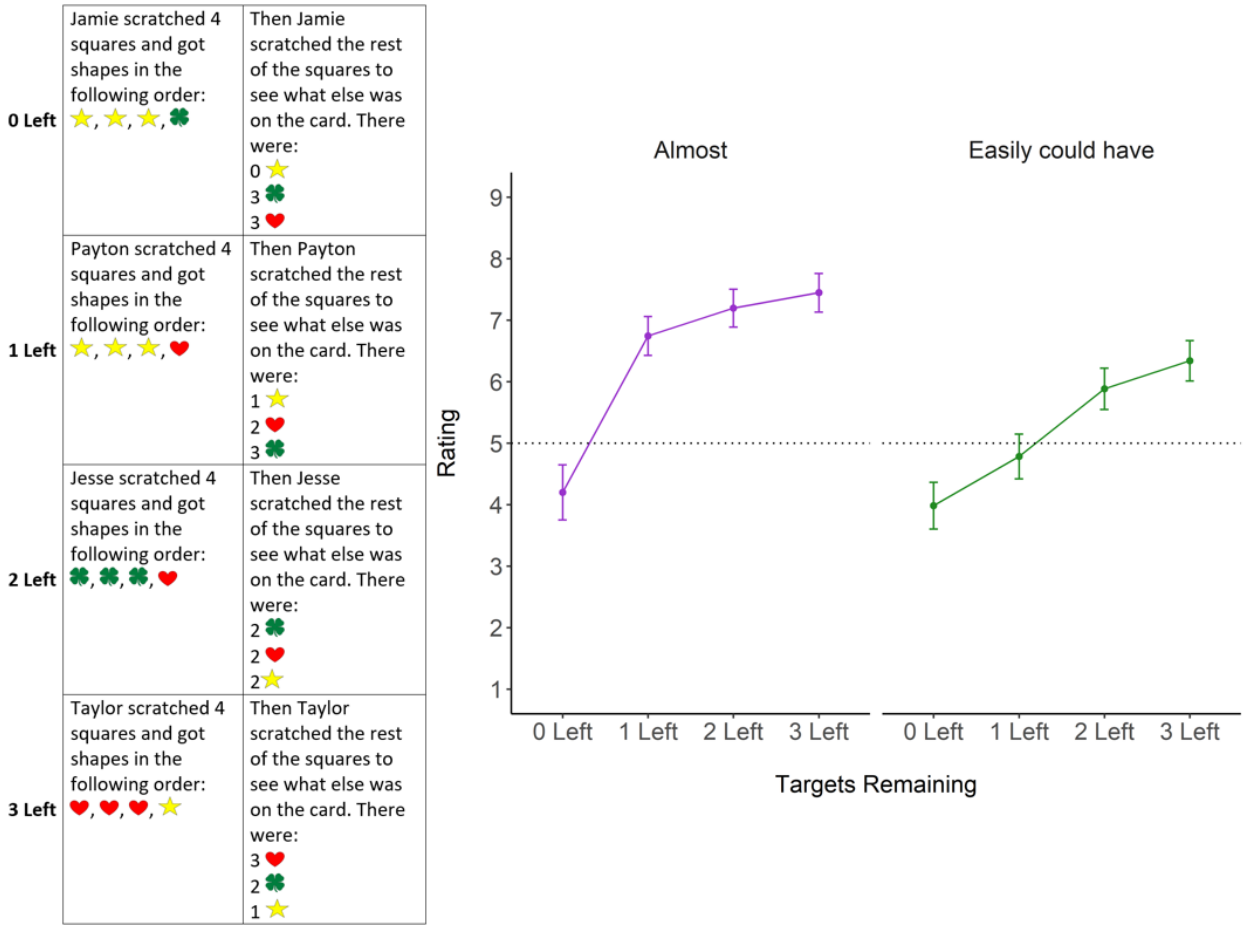

Figure 4. Stimuli of the test trials (left) and participants' responses to the close counterfactual questions (right) for Experiment 4. Vertical bars show 95% confidence intervals.

Results and Discussion

Figure 4 shows participants' mean ratings of counterfactual closeness. Ratings were higher for Almost than Easily judgments, and when more target shapes remained. Also, Almost and Easily ratings were again differentially affected by how many targets remained. Almost ratings were especially sensitive to the distinction between possible and impossible outcomes, whereas Easily ratings were not.

These observations are supported by a 4x2 GEE model with the number of targets remaining (0, 1, 2, 3) as a within-subjects factor and expression (Almost, Easily) as a between-

subjects factor. Both main effects were significant: expression (Almost>Easily), *Wald* $X^2(1) = 34.75$, $p < .001$, and targets remaining (more>less), *Wald* $X^2(3) = 303.96$, $p < .001$. There was also an interaction between expression and the number of targets remaining, *Wald* $X^2(3) = 46.58$, $p < .001$.

The interaction resulted because Almost ratings were more sensitive than Easily ratings to the distinction between 0 and 1 target remaining (Almost, $M_{difference} = 2.54$, $p < .001$, CI95% = [1.87, 3.21]; Easily, $M_{difference} = 0.80$, $p < .001$, CI95% = [0.26, 1.34]), whereas Easily ratings were more sensitive to both the distinction between 1 and 2 remaining (Easily, $M_{difference} = 1.10$, $p < .001$, CI95% = [0.66, 1.54]; Almost, $M_{difference} = 0.45$, $p = .016$, CI95% = [0.04, 0.86]) and the distinction between 2 and 3 remaining (Easily, $M_{difference} = 0.46$, $p = .017$, CI95% = [0.04, 0.87]; Almost, $M_{difference} = 0.25$, $p > .999$, CI95% = [-0.14, 0.64]).

Also, focusing on Almost judgments, ratings showed a greater increase when moving from 0 to 1 target remaining than when moving from 1 to 2 remaining, $p < .001$, CI95% = [1.57, 2.61]. But there was no significant increase between moving from 1 to 2 remaining, and moving from 2 to 3 remaining, $p = .336$, CI95% = [-0.21, 0.61]. By contrast, for Easily judgments, the increase in moving from 0 to 1 target remaining did not significantly differ from the increase in moving from 1 to 2 remaining, $p = .225$, CI95% = [-0.19, 0.79], though the increase in moving from 1 to 2 remaining was larger than the increase when moving from 2 to 3 remaining, $p = .001$, CI95% = [0.26, 1.03].

These findings again suggest that Almost judgments show extra sensitivity to the distinction between impossible and possible outcomes (i.e., over and above the corresponding increase in probability). Easily judgments instead are sensitive to increases in probability.

General Discussion

We investigated how possibility matters for judgments of counterfactual closeness. For example, we wanted to see if outcomes only seem counterfactually close if they were possible and could have happened. Our findings suggest this relation depends on the kind of close counterfactual people consider—it depends on whether they consider if an outcome *almost* happened or *easily could have*. We also investigated how propensity affects judgments of counterfactual closeness and found it had comparable effects on both kinds of closeness.

Possibility was critical for Almost judgments but not for Easily ones. With Almost ratings, participants were sharply sensitive to the distinction between possible and impossible outcomes. These ratings were much higher when the contestant had slim prior odds than when winning was impossible (Experiments 1, 3, and 4), but the ratings did not substantially change with further increases in prior odds (Experiments 2, 3, and 4). With Easily ratings, however, participants treated the distinction between possible and impossible like other equivalent differences in prior probability (Experiments 3 and 4). So, while possible and impossible alternatives were treated as qualitatively different in Almost judgments, they were viewed as only differing by degree (i.e., similar to other variations of prior probability) in Easily judgments.

Together, these findings join recent work in suggesting that there are at least two kinds of close counterfactuals. Until recently, most work on close counterfactuals focused on judgments of whether outcomes almost happened (Beck & Guthrie, 2011; Gerstenberg & Tenenbaum, 2016; Kahneman & Varey, 1990). But people also express closeness by saying outcomes easily could have happened (Levy, 2015; Teigen, 1996). A recent investigation found these two kinds of close counterfactuals are differentially affected by prior probability and the perceived distance between an outcome and a salient counterfactual alternative (Doan et al., 2023). Our current findings are consistent with the probability finding, as we found that Easily judgments were

largely determined by the prior odds of success. Moreover, our findings extend beyond this work to show that possibility more strongly affects Almost judgments.

At the same time, our findings do not directly fit with the propensity account of counterfactual closeness. On this account, people view an unrealized outcome as close when prior events seemed to be leading towards it (Kahneman & Varey, 1990). Although our findings do not contradict this account, propensity did not strongly impact judgments of closeness—its effect was small in comparison with the effects of possibility and prior probability. Also, although the propensity account predicts that impossible outcomes should never be viewed as almost happening (i.e., since prior events cannot give the impression of leading to an impossible outcome), ratings were never at floor—participants did not strongly deny that impossible outcomes almost happened. These findings might change, though, with different methods: Propensity might have greater impact if participants saw a dynamic animation rather than a static picture (see Roese et al., 2006); and ratings for impossible outcomes might be lower if participants knew which outcomes were impossible from the outset (rather than learning this after-the-fact).

Although we suggested that Almost and Easily judgments are differentially affected by possibility and probability, an alternative approach might attempt to characterize the results in terms of probabilities alone. Investigations of people's preferences between different gambles show that changes in probability are valued more when they fall close to the endpoints of absolute certainty (100%) and impossibility (0%) (for an overview, see Wu et al., 2004). For instance, improving the odds of winning a prize from 5% to 10% is seen as a greater improvement than an improvement from 30% to 35%. Applied to judgments of counterfactual closeness, future research could attempt to produce more fine-grained mappings of Almost and

Easily ratings, and this might produce a more nuanced picture of how these ratings differ from one another.

Another direction for future research might be to explore counterfactual ratings in first-person tasks – for example, where participants fill out scratch cards themselves and personally experience losses stemming from bad luck. This approach could yield different results. When people think of how events could have gone better for a fictional character, they mostly consider how the character's decisions and actions could have differed. However, when people experience or directly observe negative outcomes, they instead emphasize how the situation could have differed (Girotto et al., 2007; Pighin et al., 2011; also see Khoudary et al., 2022 for further differences between decisional and situational counterfactuals). How might this apply in our experiments? Our Easily findings fit with participants considering decisional counterfactuals, as ratings were higher when winning would have been relatively likely if the agent had decided to scratch a different square on their card. But results might differ if participants experience losses firsthand and focus on situational counterfactuals (i.e., ways the situation could have been different). Now some participants might insist they easily could have won regardless of the prior odds offered by their card. For example, they might feel winning would have been likely if they had received a different card or if the contest had different rules (e.g., if winning required getting only three matching shapes, rather than four).

A related direction for future work is investigating implications of our findings for gamblers. When people experience a near-miss while playing with scratch cards (uncover two out of three symbols needed to win), they experience higher arousal, more negative affect, and a higher urge to gamble (Stange & Dixon, 2020; Stange et al., 2016, 2017). This is because these near-miss instances signal that the player was close to winning (i.e., the propensity to win was

high). However, in many scratch card games, winning is extremely unlikely and even impossible. Our findings might suggest a way of counteracting the negative outcomes associated with near misses. Specifically, these affective outcomes might be reduced if people could be made aware of the improbability or impossibility of winning. A further avenue for future research may be to measure people's physiological responses when playing with scratch cards that have similar properties to ones we described in our experiments. This might show that people's arousal and affect differ on the basis of whether winning was possible.

Open Practices Statement

The preregistrations, stimuli, and data from all experiments are available on the Open Science Framework, https://osf.io/vyd7s/?view_only=671f7b63a1a144c68e231f66af7a456b.

References

Beck, S. R., & Guthrie, C. (2011). Almost thinking counterfactually: Children's understanding of close counterfactuals. *Child Development*, *82*, 1189-1198. https://doi.org/10.1111/j.1467-8624.2011.01590.x

Covey, J., & Zhang, Q. (2014). The effect of dynamic proximity cues on counterfactual plausibility. *Judgment and Decision Making*, *9*, 586-592.

De Brigard, F., Henne, P., & Stanley, M. L. (2021). Perceived similarity of imagined possible worlds affects judgments of counterfactual plausibility. *Cognition*, *209*, 104574. https://doi.org/10.1016/j.cognition.2020.104574

Doan, T., Denison, S., & Friedman, O. (2023). Two kinds of counterfactual closeness. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001369

Doan, T., Friedman, O., & Denison, S. (2021). Oh… So close! Children's close counterfactual reasoning and emotion inferences. *Developmental Psychology, 57*, 678-688. https://doi.org/10.1037/dev0001174

Girotto, V., Ferrante, D., Pighin, S., & Gonzalez, M. (2007). Postdecisional counterfactual thinking by actors and readers. *Psychological Science*, *18*(6), 510-515. https://doi/10.1111/j.1467-9280.2007.01931.x

Goulding, B. W. & Friedman, O. (in press). Perceived similarity explains beliefs about possibility. *Journal of Experimental Psychology: General*.

Gerstenberg, T., & Tenenbaum, J. (2016). Understanding "almost": Empirical and computational studies of near misses in Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J. (Eds.). *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 2777-2782). Cognitive Science Society.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208). Cambridge University Press.

Kahneman, D., & Varey, C. A. (1990). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology, 59*, 1101-1110. https://doi.org/10.1037/0022-3514.59.6.1101

Khoudary, A., O'Neill, K., Faul, L., Murray, S., Smallman, R., & De Brigard, F. (2022). Neural differences between internal and external episodic counterfactual thoughts. *Philosophical Transactions of the Royal Society B*, *377*(1866), 20210337. https://doi.org/10.1098/rstb.2021.0337

Klusowski, J., Small, D. A., & Simmons, J. P. (2021). Does choice cause an illusion of control?. *Psychological Science*, *32*, 159-172. https://doi.org/10.1177/0956797620958009

Kraft, J. (2012). *The epistemology of religious disagreement*. Palgrave Macmillan, New York. https://doi.org/10.1057/9781137015105_3

Kühberger, A., Großbichler, C., & Wimmer, A. (2011). Counterfactual closeness and predicted affect. *Thinking & Reasoning*, *17*, 137-155. https://doi.org/10.1080/13546783.2011.562079

Levy, J. S. (2015). Counterfactuals, causal inference, and historical analysis. *Security Studies*, *24*, 378-402. https://doi.org/10.1080/09636412.2015.1070602

Malter, M. S., Kim, S. S., & Metcalfe, J. (2021). Feelings of culpability: Just following orders versus making the decision oneself. *Psychological Science*, *32*, 635-645. https://doi.org/10.1177/09567976211002821

Markman, K. D., & Tetlock, P. E. (2000). Accountability and close-call counterfactuals: The loser who nearly won and the winner who nearly lost. *Personality and Social Psychology Bulletin*, *26*, 1213-1224. https://doi.org/10.1177/0146167200262004

Medvec, V. H., & Savitsky, K. (1997). When doing better means feeling worse: The effects of categorical cutoff points on counterfactual thinking and satisfaction. *Journal of Personality and Social Psychology*, *72*, 1284-1296. https://doi.org/10.1037/0022-3514.72.6.1284

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*, 423-429. https://doi.org/10.1111/j.1467-9280.1997.tb00455.x

Meyers-Levy, J., & Maheswaran, D. (1992). When timing matters: The influence of temporal distance on consumers' affective and persuasive responses. *Journal of Consumer Research*, *19*, 424-433. https://doi.org/10.1086/209312

Odds of a hole in one, albatross, condor and golf's other unlikely shots. (2021, March 19). *PGA*. https://www.pga.com/story/odds-of-a-hole-in-one-albatross-condor-and-golfs-unlikely-shots

Ong, D., Asaba, M., Lim, H. Y., Chen, P., & Gweon, H. (2021). "If only Santa had one more present": Exploring the development of near-miss counterfactual reasoning. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society* (pp. 2363-2369). Cognitive Science Society.

Ong, D., Goodman, N. D., & Zaki, J. (2015). Near-misses sting even when they are uncontrollable. In D. C. Noelle, R. Dale, A. D. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1775-1780). Cognitive Science Society.

Ong, D. C., Goodman, N. D., & Zaki, J. (2018). Happier than thou? A self-enhancement bias in emotion attribution. *Emotion*, *18*, 116-126. https://doi.org/10.1037/emo0000309

Phillips, J., & Cushman, F. (2017). Morality constrains the default representation of what is possible. *Proceedings of the National Academy of Sciences*, *114*(18), 4649-4654. https://doi.org/10.1073/pnas.1619717114

Pighin, S., Byrne, R. M. J., Ferrante, D., Gonzalez, M., & Girotto, V. (2011). Counterfactual thoughts about experienced, observed, and narrated events. *Thinking & Reasoning, 17*(2), 197–211. https://doi.org/10.1080/13546783.2011.561598

Roese, N. J., Fessel, F., Summerville, A., Kruger, J., & Dilich, M. A. (2006). The propensity effect: When foresight trumps hindsight. *Psychological Science*, *17*, 305-310. https://doi.org/10.1111/j.1467-9280.2006.01703.x

Shtulman, A. (2009). The development of possibility judgment within and across domains. *Cognitive Development*, *24*(3), 293-309. https://doi.org/10.1016/j.cogdev.2008.12.006

Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development*, *78*(3), 1015-1032. https://doi.org/10.1111/j.1467-8624.2007.01047.x

Stange, M., & Dixon, M. J. (2020). Scratch card near-miss outcomes increase the urge to gamble, but do not impact further gambling behaviour: A pre-registered replication and extension. *Journal of Gambling Studies*, *36*, 887-902. https://doi.org/10.1007/s10899-020-09932-7

Stange, M., Grau, M., Osazuwa, S., Graydon, C., & Dixon, M. J. (2017). Reinforcing small wins and frustrating near-misses: Further investigation into scratch card gambling. *Journal of Gambling Studies*, *33*, 47-63. https://doi.org/10.1007/s10899-016-9611-0

Stange, M., Graydon, C., & Dixon, M. J. (2016). “I was that close”: Investigating players’ reactions to losses, wins, and near-misses on scratch cards. *Journal of Gambling Studies*, *32*, 187-203. https://doi.org/10.1007/s10899-015-9538-x

Sweeny, K., & Vohs, K. D. (2012). On near misses and completed tasks: The nature of relief. *Psychological Science*, *23*, 464-468. https://doi.org/10.1177/0956797611434590

Teigen, K. H. (1996). Luck: The art of a near miss. *Scandinavian Journal of Psychology*, *37*, 156-171. https://doi.org/10.1111/j.1467-9450.1996.tb00648.x

Wu, G. Zhang, J., & Gonzalez, R. (2014). Decision under risk. In J. Koehler & N. Harvey (Eds.,) *Blackwell handbook of judgment and decision making* (pp. 399-423). Blackwell.